

Application for
UNITED STATES LETTERS PATENT

of

YOSHU YOSHIBA
KAZUKO SHINOZAKI

and

KAZUO SHINOZAKI

for

**TRANSGENIC RICE PLANT AND ITS FAMILY WITH
ENVIRONMENTAL STRESS RESISTANT BY PROLINE
ACCUMULATION OF HIGH LEVEL AND ITS PRODUCTION**

10026767-122701

SPECIFICATION

TITLE OF THE INVENTION

Transgenic rice plant and its family with environmental
stress resistant by proline accumulation of high level
and its production

BACKGROUND OF THE INVENTION

The present invention relates to a rice plant
having a high level of proline accumulating ability,
and improved salinity-tolerance, drought-tolerance, and
low temperature-tolerance, and its production method.

It is known that, for several plants including
halophytes, when the plants are subjected to a high
salinity stress or a drought stress, they accumulate
proline, which is one of amino acids, in their
cytoplasm. This is considered useful for regulating
the osmotic pressure in the plant cytoplasm, or
inhibiting the degradation of a functional protein due
to the stress. The proline in a plant is synthesized
from a glutamic acid by two enzymes of a Δ^1 -pyrroline-
5-carboxylate (P5C) synthetase (P5CS) and a P5C
reductase. On the other hand, proline is degraded into
a glutamic acid by the two enzymes of a proline
dehydrogenase (ProDH) and a P5C dehydrogenase.

When each of the aforesaid plants is subjected
to a water stress (the state in which water is
difficult to absorb) such as a high salinity stress or
a drought stress, the expression level of the P5CS gene

10025767-122704
is increased to activate the P5CS. However, the P5CR
activity and the gene expression are constant at a low
level. Further, the gene expression and the enzyme
activity related to metabolism are also in the
5 inhibited states. However, once the water stress has
been removed, conversely, this time, the gene
expression and enzyme activity related to biosynthesis
are inhibited, so that the expression of the ProDH gene
is rapidly induced, and the enzyme activity is also
10 enhanced. As a result, the proline accumulated in the
cytoplasm is rapidly metabolized to a glutamic acid.

15 From the foregoing description, it is
considered that the P5CS becomes rate-limiting for
proline synthesis under a water stress. Whereas, the
ProDH becomes rate-limiting for proline metabolism
after releasing the water stress (Yoshida et al., Plant
Cell Physiol, 38: 1095 - 1102 (1997)).

SUMMARY OF THE INVENTION

20 It is predicted that food shortage due to an
expansion of the saline soil area caused by drought and
semi-drought with the deterioration of global
environment, and population growth will become
increasingly more serious in the future. Researches
25 have been pursued in diversified fields respectively on
the breeding of crop plants resistant to a high
salinity stress, a drought stress, and a low
temperature stress (the state in which water is

difficult to absorb) as those playing an important role in solving the world food problem, and the results are expected to be promising.

It is an object of the present invention to provide a rice plant which has a high proline accumulating ability, and accordingly has improved salinity-tolerance, drought-tolerance, and low temperature-tolerance by focusing attention on the importances of a Δ^1 -pyrroline-5-carboxylate (P5C) synthetase (P5CS) and a proline dehydrogenase (ProDH) which are the rate-limiting enzymes related to synthesis and metabolism of proline in plants, and regulating the expression of genes for the enzymes with a gene recombination technology, and its production method.

The P5CS gene related to proline synthesis is introduced to be overexpressed; the antisense (reverse DNA sequence-containing) gene of the ProDH gene related to the metabolism is introduced to inhibit the degradation of proline; or both the P5CS gene and the antisense gene of the ProDH gene are introduced to promote the proline synthesis while inhibiting the degradation of proline. As a result, proline is accumulated with a high concentration in the cells of rice and a rice plant.

In the present invention, by accumulation of proline at a high concentration, it becomes possible to perform molecular breeding of rice and a rice plant

having salinity-tolerance, drought-tolerance, or low temperature-tolerance.

Heretofore, there is known no report that an increase in concentration of proline as an osmoprotectant is allowed by synthesis promotion and degradation inhibition in rice and a rice plant. The inventors of the present invention have focused attention on the importances of the P5CS gene and the ProDH gene. Then, in order to solve novel technical problems which have not been known in the prior art, they have conducted studies from various fields including the study on the selection of the rice variety into which the gene is easily introduced, the study for improving the callus formation rate, the study on the construction of a vector for introducing the gene for rice, and the like. In consequence, they have provided novel technical elucidation, resulting in the completion of the present invention.

In the present invention, there are provided a rice plant transformed by introducing therein the proline synthesis gene and the antisense gene of the proline metabolism gene derived from rice or *Arabidopsis thaliana* individually or in combination, and its production method.

In the rice plant of the present invention, either or both of the gene encoding the synthetase protein of proline which is one of amino acids and the antisense gene of the proline dehydrogenase have been

introduced. With this construction, it is possible to implement a rice plant having improved salinity-tolerance, drought-tolerance, and low temperature-tolerance. Further, the mature rice seeds gathered from the rice plant of the present invention, particularly the rice seeds are characterized by keeping a high proline accumulating ability over a plurality of generations.

Further, the present invention is targeted for rice and rice plants. The targets have no particular restriction as long as they are the plants belonging to the rice plants. Examples of the plants belonging to the rice plants include rice, corn, wheat, barley, rye, turf, millet, and barn grass. In particular, the present invention can be more preferably applied to rice.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A to 1D are diagrams respectively showing the vectors for rice in which proline synthesis-related enzyme P5CS genes and proline metabolism-related enzyme ProDH genes, and antisense genes thereof have been respectively incorporated;

FIG. 2 is a graph showing the amount of proline accumulated in rice lines under no stress in which the vectors shown in FIGS. 1A to 1D have been respectively introduced by genetic engineering; and

FIG. 3 is a graph showing the salinity-

tolerance of each of the transgenic rice lines in which the proline-related genes have been respectively incorporated shown in FIG. 2.

5 DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In rice plants of examples of the present invention, either or both of the proline (osmoprotectant) synthesis gene and the antisense gene of the proline metabolism derived from rice or 10 Arabidopsis thaliana gene have been introduced for transformation.

Examples of one type of gene to be introduced to the rice plants of the examples of the present invention include: (1) a P5CS (Δ^1 -pyrroline-5-carboxylate (P5C) synthetase) gene of rice containing the sequence (DNA sequence and amino acid sequence) 15 according to SEQ ID No. 1; (2) a P5CS (Δ^1 -pyrroline-5-carboxylate (P5C) synthetase) gene of Arabidopsis thaliana containing the sequence (DNA sequence and amino acid sequence) according to SEQ ID N2; and (3) 20 the antisense (reverse DNA sequence-containing) gene of the ProDH (proline dehydrogenase) gene of Arabidopsis thaliana containing the sequence (DNA sequence and amino acid sequence) according to Seq ID NO. 3.

25 Examples of the two types of genes to be introduced into the rice plants of the examples of the present invention include:

(1) Two genes of the P5CS (Δ^1 -pyrroline-5-carboxylate

10026767.122701

(P5C) synthetase) of rice containing the sequence according to SEQ ID NO. 1 or the P5CS gene of Arabidopsis thaliana containing the sequence according to SEQ ID NO. 2, and the antisense (reverse DNA sequence-containing) gene of the ProDH (proline dehydrogenase) gene of Arabidopsis thaliana containing the sequence according to SEQ ID NO. 3; and

(2) Tandemly connected two genes of the P5CS (Δ^1 -pyrroline-5-carboxylate (P5C) synthetase) gene of rice containing the sequence according to SEQ ID NO. 1 or the P5CS gene of Arabidopsis thaliana containing the sequence according to SEQ ID NO. 2, and the antisense (reverse DNA sequence-containing) gene of the ProDH (proline dehydrogenase) gene of Arabidopsis thaliana containing the sequence according to SEQ ID NO. 3.

In each of the vectors to be used in the examples of the present invention, there is incorporated any one gene of the P5CS (Δ^1 -pyrroline-5-carboxylate (P5C) synthetase) gene of rice containing the sequence according to SEQ ID NO. 1, the P5CS gene of Arabidopsis thaliana containing the sequence according to SEQ ID NO. 2, and the antisense (reverse DNA sequence-containing) gene of the ProDH (proline dehydrogenase) gene of Arabidopsis thaliana containing the sequence according to SEQ ID NO. 3. Alternatively, there are incorporated two genes of the P5CS gene of rice or Arabidopsis thaliana, and the aforesaid antisense gene in tandemly connected relation to each

other.

The rice plants of the examples of the present invention can be obtained by, for example, any of the following methods.

- 5 (1) The aforesaid vector is introduced into the calli derived from a rice plant, and the calli are grown. Then, a plant body is regenerated from the calli;
- 10 (2) The aforesaid vector is introduced into the protoplast derived from a rice plant, and a plant body is regenerated from the colony obtained by growing the protoplast; and
- (3) Crossing with the rice plants obtained by introducing the vector therein by genetic engineering is carried out.

15 Examples of the production method of the rice plants of the examples of the present invention include the following methods:

- 20 (1) The aforesaid vector is introduced into the calli derived from a rice plant by using *Agrobacterium tumefaciens*, and the calli are grown. Then, a plant body is regenerated from the calli;
- (2) The aforesaid vector is introduced into the protoplast derived from a rice plant by electroporation, and a plant body is regenerated from the colony
- 25 obtained by growing the protoplast; and
- (3) Crossing with the rice plants obtained by introducing the vector therein by genetic engineering is carried out.

These production methods provide a rice plant having a high proline accumulating ability, and having improved salinity-tolerance, drought-tolerance, and low temperature-tolerance levels.

5 Further, mature seeds gathered from the rice plants of the examples of the present invention, particularly the rice seeds will maintain their high proline accumulating abilities over a plurality of generations.

10 The rice plants of the examples of the present invention and its production method will be described in details by way of embodiments thereof by using rice as a typical example step by step below. It is
15 needless to say that the steps described below are applicable to other rice plants than rice with or without changing the various conditions.

(Gene cloning)

First, a mRNA is extracted from a rice seedling. A cDNA is synthesized by using the mRNA. The cDNA is
20 combined with a vector made of a plasmid or a phage, and introduced into E. coli to prepare a recombinant DNA. The resulting transformant in which the recombinant DNA has been introduced is subjected to
25 screening by plaque hybridization using the P5CS gene from Arabidopsis thaliana as a probe. The sequences of the P5CS genes from rice and Arabidopsis thaliana have been already reported (Yoshiba et al., Plant J. (1995) 7:751-760, and Igarashi et al., Plant Mol. Biol. (1997)

33:857-865). Based on these reports, appropriate primers are designed, and subjected to screening by PCR to select a target transformant. A target plasmid is isolated from the transformant obtained. If required, it is cut with an appropriate restriction enzyme, and subjected to subcloning in a plasmid vector for cloning. It is also possible to subject the P5CS gene of *Arabidopsis thaliana* to cloning in the same manner as with rice. However, as a sample from which a mRNA is to be extracted, the one subjected to a high salinity stress (immersed in a 250 mM NaCl solution or the like) or the one subjected to a drought stress treatment is more preferable than the one bred under a normal environment. This is because the P5CS gene is induced in response to a water stress such as a high salinity stress or a drought stress (Yoshida et al., *Plant J.* (1995) 7: 751-760, Igarashi et al., *Plant Mol. Biol.* (1997) 33: 857-865, and Yoshida et al., *Plant Cell Physiol.* (1997) 38: 1095-1102).

On the other hand, it is also possible to subject the ProDH gene of *Arabidopsis thaliana* (its sequence has already been reported in Kiyosue et al., *Plant Cell* (1996) 8:1323-1335) to cloning in the foregoing manner. However, as the sample from which a mRNA is to be extracted, there may be used the one which has been subjected to a drought stress (about 10-hour treatment), then immersed in water again, and allowed to absorb water, the one which has been

immersed in a proline solution, and allowed to absorb proline, or the like. This is due to the following fact. Namely, the ProDH gene is inhibited from its expression under a water stress, and the gene
5 expression is induced by a high concentration of proline (Kiyosue et al., Plant Cell (1996) 8: 1323-1335, and Yoshiba et al., Plant Cell Physiol. (1997) 38: 1095-1102).

10 If the samples as described above are used, it is possible to isolate the P5CS gene and the ProDH gene not only from rice or Arabidopsis thaliana but also from other rice plants.

(Construction of gene introduction vector)

15 Respective P5CS genes and ProDH genes subjected to cloning are cut from plasmids with appropriate restriction enzymes, and, as shown in FIGS. 1A to 1D, each is combined behind the 35S promoter of a cauliflower mosaic virus of a vector for rice obtained by modifying a pBI vector. In FIGS. 1A to 1D, RB
20 denotes the right border, 35SPro denotes the promoter of a cauliflower mosaic virus, P5CS denotes the proline synthesis-related enzyme gene of rice or Arabidopsis thaliana, ProDH denotes proline metabolism-related enzyme gene of Arabidopsis thaliana, Noster denotes the
25 terminator of a nopaline synthetase gene, HTP denotes a hygromycin resistant gene, and LB denotes the left border. Whereas, each of the arrows indicates the orientation of the sense of each gene.

In FIGS. 1A to 1D, FIG. 1A is a diagram showing an example of the vector (construct) so constructed that the sequence in the order of RB-35SPro-P5CS-Noster-35SPro-HTP-Noster-LB has been achieved. FIG. 1B is a diagram showing an example in which, with respect to FIG. 1A, the same sequence in the order of RB-35SPro-P5CS-Noster-35SPro-HTP-Noster-LB as in the construct of FIG. 1A has been achieved, but the gene P5CS has been sequenced in antisense orientation. FIG. 1C is a diagram showing an example in which the gene ProDH has been sequenced in antisense orientation, and substituted for the gene P5CS of the construct of FIG. 1A, to construct a vector with a sequence in the order of RB-35SPro-ProDH (antisense)-Noster-35SPro-HTP-Noster-LB. FIG. 1D is a diagram showing an example in which, to the construct of FIG. 1A, the gene ProDH has been further sequenced in antisense orientation, and the construct shown in FIG. 1C has been further connected thereto in tandem, to construct a vector with a sequence in the order of RB-35SPro-P5CS-Noster-35SPro-ProDH (antisense)-Noster-35SPro-HTP-Noster-LB.

The 35S promoter is well known as a promoter which is strong and invariably induces the gene expression in any tissue. As for the orientation in which the gene is incorporated, the P5CS gene is connected in the sense orientation, and the ProDH gene in the antisense orientation.

Then, each vector to which each of the genes

has been connected is introduced into *Agrobacterium tumefaciens* EHA 101 by electroporation. The *Agrobacterium tumefaciens* in which each construct (FIGS 1A to 1D) has been introduced is cultured and grown in a YEP medium containing Bacto Pepton (10 g/l), Bacto Yeast Extract (10 g/l), sodium chloride (5 g/l), 1M magnesium chloride (2 ml/l), and hygromycin B (50 mg/l) at 28 °C. Gene introduction is carried out by infecting the callus cell of rice with the *Agrobacterium tumefaciens* into which each construct (FIGS. 1A - 1D) has been introduced. The construct D is so designed that the two genes (the P5CS gene and the ProDH gene) are connected to each other in tandem to be simultaneously introduced. However, even if the constructs A and C are mixed for coinfection, it is also possible to obtain the same effects as with the construct D.

Incidentally, a HPT (hygromycin resistant) gene is connected to each construct. This is for efficiently selecting the cell and plant body transformed for the basic research on analysis of the effects of the introduced genes. Therefore, the HPT gene is not required to be incorporated therein for actual cultivation on the salt damaged land or the dry land.

(Induction of rice calli for gene introduction)

Mature rice seeds are sterilized with 70 % ethyl alcohol for 10 minutes, and with 3 % sodium

10065767-122701
hypo-chlorite for 1 hour after stripping the hulls
therefrom. After sterilization, the seeds are washed
with sterilized water 3 times, and bedded on a pH 5.8
N6 medium (2N6 medium) containing 1 g/l casamino acid,
5 30 g/l sucrose, 2 mg/l 2,4-dichlorophenoxyacetic acid,
and 2 g/l Gelrite, and cultured at 28 °C in the dark
for 3 to 5 weeks.

(Gene introduction into rice calli)

10 Out of the rice calli induced in the foregoing
manner, the ones with a size of 1 to 3 mm are bedded on
the 2N6 medium again, and cultured at 28 °C in the dark
for 3 to 4 days. As a result, it is possible to
enhance the division activity of the callus cell. The
gene introduction is carried out by mixing the cultured
15 calli and a solution of each construct-introduced
Agrobacterium tumefaciens grown in the YEP medium (the
solution diluted so that the concentration of the
bacteria is 0.1 as determined at OD 660nm) for
infection. Thereafter, the calli are cultured at 25 °C
20 in the dark for 3 days. After cultivation, the calli
are washed and sterilized several times by a cefotaxime
aqueous solution with a concentration of 1 mg/4 ml to
remove extra bacteria attached to the surfaces of the
calli, and cleaned with a sterilized kim towel or the
25 like. Subsequently, it is bedded on a 2N6 medium
(secondary selection medium) containing 250 mg/l
cefotaxime and 10 mg/l hygromycin B, and cultured at
28 °C in the dark for 1 week.

(Selection of transformed calli and
regeneration of plant body)

10026767.122701

The calli cultured in the medium containing
cefotaxime is bedded on a medium (secondary selection
5 medium) in which the content of hygromycine B has been
increased to 30 mg/l, and cultured at 28 °C in the dark
for 3 weeks. Thereafter, the calli are transferred to
a pH 5.8 MS medium (regeneration induction medium)
containing 30 g/l sucrose, 30 g/l sorbitol, 2 g/l
10 casamino acid, 11 g/l MES buffer, 2 mg/l NAA, 1 mg/l
kinetin, 250 mg/l cefotaxime, 30 mg/l hygromycine B,
and 4 g/l Gelrite, and cultured in the bright place at
28 °C for 3 week. The gene-introduced calli form a
green spot, from which shoots and roots are regenerated.
15 The regenerated calli are further transferred to a pH
5.8 MS medium (plant body formation medium) containing
30 g/l sucrose, 250 mg/l cefotaxime, 30 mg/l
hygromycine B, and 8 g/l agar, from which plant
hormones have been removed, and cultured in the bright
20 place at 28 °C for several weeks. In consequence, the
plant body is bred more largely.

(Breeding of transformed rice plant body and
seed formation)

Upon having grown to a seedling height of about
25 4 to 5 cm in a petri dish, the regenerated rice is
transferred to a planter in which the soil for raising
seedling is placed. Then, it is bred in an artificial
climate system with an illuminance of about 20,000 lx

under a temperature condition of 28 °C until the fourth leaf to the fifth leaf develop. Subsequently, the seedling is further transferred into a pot containing the soil into which a fertilizer has been appropriately added, and bred in a greenhouse until the seeds ripen. Assuming that the present generation of the plant body regenerated is of the T0 generation, and that the seeds obtainable from this plant body is of the T1 generation, the ones of the T2 to T3 generations are bred. When they are cultivated in an actual farm land, they are required to be commercialized after carrying out the various safety evaluation tests over further generations, and confirming the safety.

(Extraction of proline from transformed rice and concentration measurement thereof)

Proline is extracted from the leaves of the seedling (whose forth leaf has developed) of the transformed rice of the T2 generation or the T3 generation. The leaves of the rice seedling bred in the artificial climate system are cut off in an amount of about 200 mg by scissors or the like. Then, in a mortar, liquid nitrogen is added thereto, and the leaves are ground into powder. The resulting sample in powder form is mixed with pure water, and further milled by means of a homogenizer or the like. The milled sample is heated at 97 °C for 6 minutes, and then ice cooled. The sample is then centrifuged at about 17,000 ×G for 10 minutes at 4 °C to separate the

supernatant. To the supernatant obtained, a trichloroacetic acid is added and mixed so that the final concentration is 5 %. The resulting mixture is then centrifuged at about 17,000 XG for 10 minutes at 4 °C again to precipitate protein. Proline as an osmoprotectant is contained in the supernatant at this step, and the concentration thereof is determined by means of high performance liquid chromatography (HPLC). The qualitative determination of proline is carried out in the following manner. The solutions in which various amino acids have been dissolved to a given concentration are previously determined by HPLC. The amount of proline contained in the leaf of an actual transgenic rice is determined based on the retention times.

FIG. 2 shows the proline content of each of the transgenic rice lines under no stress into which various genes have been introduced. The hollow graphs in the leftmost column represent control samples into which proline-related genes have not been incorporated. Whereas, the solidly shaded graphs in the right-hand five columns denote respective transgenic rice lines into which proline-related genes have been incorporated. It is indicated that the proline content varies according to the type of the gene introduced.

There is observed almost no accumulation for each sample in which the P5CS gene (OsP5CS) of rice has been introduced in antisense orientation (FIG. 1B) in

the second column from left. For each sample in which the P5CS gene (AtP5CS) of *Arabidopsis thaliana* has been introduced in sense orientation (FIG. 1A) in the third column from left, there is observed an increase in amount of proline accumulated over the control samples. Similarly, for each sample in which the ProDH gene (AtProDH) of *Arabidopsis thaliana* has been introduced in antisense orientation (FIG. 1C) and each sample in which the P5CS gene (OsP5CS) of rice has been introduced in sense orientation (FIG. 1A) in the fourth and fifth columns from left, respectively, there are observed increases in amount of proline accumulated over the control sample. In contrast to these, for each sample in which the P5CS gene (OsP5CS) of rice has been introduced in sense orientation, and the ProDH gene (AtProDH) of *Arabidopsis thaliana* in antisense orientation in the rightmost column, there is observed a considerably larger amount of proline accumulated (100 times or more with respect to the control sample for the case where the amount of proline accumulated is larger) as compared with each of the aforesaid samples in which one type of gene has been introduced. Then, it is indicated that each sample of OsP5CS (in the fifth column from left) is slightly more effective for proline accumulation than each sample of AtP5CS (in the third column from left) among the samples in which genes have been introduced in sense orientation.

(Salinity tolerance test and improvement of

salinity tolerance of transgenic rice)

FIG. 3 shows the results of a salinity tolerance test performed at a 250 mM concentration (about half the salt concentration of sea water) by using several lines of the transgenic rice for which proline accumulation has been observed shown in the right hand four columns of FIG. 2. The hollow graphs denote the control samples in which proline related genes have not been incorporated. Whereas, the solidly shaded graphs denote the transgenic rice samples. The salinity tolerance test was carried out in accordance with the testing method using known survival rates as indexes (Japanese Published Unexamined Patent Application No. Hei 09-266726, title of the invention: evaluation of salt resistance of plant). It has been shown that the control samples in which proline-related genes have not been introduced die 5 days after a salt treatment, while the transgenic rice samples which accumulate proline show high survival rates, i.e., 95 % for the third day, and 65 % even after the five-day treatment. This indicates that the salinity tolerance can be improved by transforming rice, and thereby enhancing the proline accumulating ability thereof.

Therefore, if the gramineous crop produced according to the present invention is subjected to breeding by further pursuing detailed analysis such as the safety evaluation thereon, it becomes capable of being cultured in the salt accumulated soil or the

desertified soil. Therefore, food productivity can be expected to be improved. Further, it can be largely expected that the crop plant is also capable of coping with the population growth in developing countries.

- 5 In accordance with the present invention, it has become possible to produce a transgenic rice plant having an enhance proline accumulating ability. Further, for the rice plant produced by the method of the present invention, the amount of proline
- 10 accumulated therein has been increased, so that it has become possible to improve the salinity tolerance level thereof.

[Sequence Listing]

<110> Hitachi, LTD.

RIKEN

Japan International Research Center for
Agricultural Science

Bio-oriented Technology Research
Advancement Institute (BRAIN)

<120> Transgenic rice plant and its family with
environmental stress resistant by proline
accumulation of high level and its production.

<130> NT01P0353

<160> 3

<210> 1

<211> 2549

<212> DNA

<213> *Oryza sativa* L.

<220>

<221> CDS

<222> 99..2249

<300>

<301> Yumiko Igarashi, Yoshu Yoshiba, Yukika
Sanada, Kazuko Yamaguchi-Shinozaki, Keishiro Wada,
Kazuo Shinozaki

<302> Characterization of the gene for Δ^1 -
pyrroline-5-carboxylate synthetase and correlation
between the expression of the gene and salt
tolerance in *Oryza sativa* L.

<303> Plant Molecular biology

<304> 33

<306> 857-865

<307> 1996-12-03

<308> D49714

<309> 1995-03-16

<400> 1

gcggctgcgg cggaagcgc gcgagcgtg ggagaggat ttacaggtag agggagaggg 60

tggaggagga gaggctgagg ctaggaagcg gtttcgcc atg gcg agc gtc gac ccg 116

Met Ala Ser Val Asp Pro

1

5

tcc cgg agc ttc gtg agg gac gtg aag cgc gtc atc atc aag gtg ggc 164

Ser Arg Ser Phe Val Arg Asp Val Lys Arg Val Ile Ile Lys Val Gly

10

15

20

act gca gtt gtc tcc aga caa gat gga aga ttg gct ttg ggc agg gtt 212

Thr Ala Val Val Ser Arg Gln Asp Gly Arg Leu Ala Leu Gly Arg Val

25

30

35

gga gct ctg tgc gag cag gtt aag gaa ctg aac tct tta gga tac gaa 260

Gly Ala Leu Cys Glu Gln Val Lys Glu Leu Asn Ser Leu Gly Tyr Glu

40

45

50

gtg att ttg gtc acc tca ggt gct gtt gga gtg ggg cga cag cga ctt 308

Val Ile Leu Val Thr Ser Gly Ala Val Gly Val Gly Arg Gln Arg Leu

55

60

65

70

ttg gca ctg gaa ctg aaa gct gat ctc ctt att ctg ctc agt gat gtg	692
Leu Ala Leu Glu Leu Lys Ala Asp Leu Leu Ile Leu Leu Ser Asp Val	
185 190 195	
gat ggg ttg tat agt ggt cca cca agt gaa cca tca tca aaa atc ata	740
Asp Gly Leu Tyr Ser Gly Pro Pro Ser Glu Pro Ser Ser Lys Ile Ile	
200 205 210	
cac act tat att aaa gaa aag cat cag caa gaa atc act ttt gga gac	788
His Thr Tyr Ile Lys Glu Lys His Gln Gln Glu Ile Thr Phe Gly Asp	
215 220 225 230	
aaa tct cgt gta ggt aga gga ggc atg aca gca aaa gtg aag gct gct	836
Lys Ser Arg Val Gly Arg Gly Gly Met Thr Ala Lys Val Lys Ala Ala	
235 240 245	
gtc ttg gct tca aat agc ggc aca cct gtg gtt att aca agt ggg ttt	884
Val Leu Ala Ser Asn Ser Gly Thr Pro Val Val Ile Thr Ser Gly Phe	
250 255 260	
gaa aat cgg agc att ctt aaa gtt ctt cat ggg gaa aaa att ggt act	932
Glu Asn Arg Ser Ile Leu Lys Val Leu His Gly Glu Lys Ile Gly Thr	
265 270 275	
ctc ttt cac aag aat gcg aat ttg tgg gaa tca tct aag gat gtt agt	980
Leu Phe His Lys Asn Ala Asn Leu Trp Glu Ser Ser Lys Asp Val Ser	
280 285 290	

gtt ctc tta att gtt ttt gag tcc cga cct gat gcc ttg gtt cag att	1364
Val Leu Leu Ile Val Phe Glu Ser Arg Pro Asp Ala Leu Val Gln Ile	
410 415 420	
gca tct ttg gca att cga agt ggt aat ggt ctt ctc cta aaa ggt gga	1412
Ala Ser Leu Ala Ile Arg Ser Gly Asn Gly Leu Leu Leu Lys Gly Gly	
425 430 435	
aaa gaa gct atc aga tca aac acg ata ttg cat aag gtt ata act gat	1460
Lys Glu Ala Ile Arg Ser Asn Thr Ile Leu His Lys Val Ile Thr Asp	
440 445 450	
gct att cct cgt aat gtt ggt gaa aaa ctt att ggc ctt gtt aca act	1508
Ala Ile Pro Arg Asn Val Gly Glu Lys Leu Ile Gly Leu Val Thr Thr	
455 460 465 470	
aga gat gag atc gca gat ttg cta aag ctt gat gat gtc att gat ctt	1556
Arg Asp Glu Ile Ala Asp Leu Leu Lys Leu Asp Asp Val Ile Asp Leu	
475 480 485	
gtc act cca aga gga agt aat aag ctt gtc tct caa atc aag gcg tca	1604
Val Thr Pro Arg Gly Ser Asn Lys Leu Val Ser Gln Ile Lys Ala Ser	
490 495 500	
act aag att cct gtt ctt ggg cat gct gat ggt ata tgc cac gta tat	1652
Thr Lys Ile Pro Val Leu Gly His Ala Asp Gly Ile Cys His Val Tyr	
505 510 515	

att gac aaa tca gct gac atg gat atg gca aaa ctt att gta atg gat	1700
Ile Asp Lys Ser Ala Asp Met Asp Met Ala Lys Leu Ile Val Met Asp	
520 525 530	
 gca aaa act gat tac cca gca gcc tgc aat gca atg gag acc tta cta	1748
Ala Lys Thr Asp Tyr Pro Ala Ala Cys Asn Ala Met Glu Thr Leu Leu	
535 540 545 550	
 gtt cat aag gat ctt atg aag agt cca ggc ctt gac gac ata tta gta	1796
Val His Lys Asp Leu Met Lys Ser Pro Gly Leu Asp Asp Ile Leu Val	
555 560 565	
 gca cta aaa aca gaa gga gtt aat att tat ggt gga cct att gcg cac	1844
Ala Leu Lys Thr Glu Gly Val Asn Ile Tyr Gly Gly Pro Ile Ala His	
570 575 580	
 aaa gct ctg gga ttt cca aaa gct gtt tca ttt cat cat gag tat agt	1892
Lys Ala Leu Gly Phe Pro Lys Ala Val Ser Phe His His Glu Tyr Ser	
585 590 595	
 tct atg gcc tgc act gtt gag ttt gtt gat gat gtt caa tca gca att	1940
Ser Met Ala Cys Thr Val Glu Phe Val Asp Asp Val Gln Ser Ala Ile	
600 605 610	
 gac cat att cat cgt tat gga agt gct cat aca gat tgt atc gtc act	1988
Asp His Ile His Arg Tyr Gly Ser Ala His Thr Asp Cys Ile Val Thr	
615 620 625 630	

acacacaaaa taaattcatc agttctgaga gcaagatttt ggaggttcag cttctccatg 2456

taataagtaa attcagttct gagaacttgt gtaccaacgc gctatgttgc ttgtaatgag 2516

cgatactaac atctgtgatt gcacatatac taa 2549

<210> 2

<211> 2571

<212> DNA

<213> *Arabidopsis thaliana*

<220>

<221> CDS

<222> 107...2260

<301> Yoshu Yoshiba, Tomohiro Kiyasue, Takeshi Katagiri, Hiroko Ueda, Tsuyoshi Mizoguchi, Kazuko Yamaguchi-Shinozaki, Keishiro Wada, Yoshinori Harada, Kazuo Shinozaki

<302> Correlation between the induction of a gene for Δ^1 -pyrroline-5-carboxylate synthetase and the accumulation of proline in *Arabidopsis thaliana* under osmotic stress.

<303> The Plant Journal

<304> 7

<305> 5

<306> 751-760

<307> 1995-01-20

<308> D32138

<309> 1994-07-12

<400> 2

ctgatattta ttttcttacc ttaaatacga cggtgcttca ctgagtcga ctcagttaac 60

tcgttcctct ctcgtgtgt ggttttggta gacgacgacg acgata atg gag gag	115
Met Glu Glu	
1	
cta gat cgt tca cgt gct ttt gcc aga gac gtc aaa cgt atc gtc gtt	163
Leu Asp Arg Ser Arg Ala Phe Ala Arg Asp Val Lys Arg Ile Val Val	
5 10 15	
aag gtt ggg aca gca gtt gtt act gga aaa ggt gga aga ttg gct ctt	211
Lys Val Gly Thr Ala Val Val Thr Gly Lys Gly Gly Arg Leu Ala Leu	
20 25 30 35	
ggt cgt tta gga gca ctg tgt gaa cag ctt gcg gaa tta aac tcg gat	259
Gly Arg Leu Gly Ala Leu Cys Glu Gln Leu Ala Glu Leu Asn Ser Asp	
40 45 50	
gga ttt gag gtg ata ttg gtg tca tct ggt gcg gtt ggt ctt ggc agg	307
Gly Phe Glu Val Ile Leu Val Ser Ser Gly Ala Val Gly Leu Gly Arg	
55 60 65	
caa agg ctt cgt tat cga caa tta gtc aat agc agc ttt gcg gat ctt	355
Gln Arg Leu Arg Tyr Arg Gln Leu Val Asn Ser Ser Phe Ala Asp Leu	
70 75 80	
cag aag cct cag act gaa ctt gat ggg aag gct tgt gct ggt gtt gga	403
Gln Lys Pro Gln Thr Glu Leu Asp Gly Lys Ala Cys Ala Gly Val Gly	
85 90 95	

caa agc agt ctt atg gct tac tat gag act atg ttt gac cag ctt gat 451
 Gln Ser Ser Leu Met Ala Tyr Tyr Glu Thr Met Phe Asp Gln Leu Asp
 100 105 110 115

gtg acg gca gct caa ctt ctg gtg aat gac agt agt ttt aga gac aag 499
 Val Thr Ala Ala Gln Leu Leu Val Asn Asp Ser Ser Phe Arg Asp Lys
 120 125 130

gat ttc agg aag caa ctt aat gaa act gtc aag tct atg ctt gat ttg 547
 Asp Phe Arg Lys Gln Leu Asn Glu Thr Val Lys Ser Met Leu Asp Leu
 135 140 145

agg gtt att cca att ttc aat gag aat gat gct att agc acc cga aga 595
 Arg Val Ile Pro Ile Phe Asn Glu Asn Asp Ala Ile Ser Thr Arg Arg
 150 155 160

gcc cca tat cag gat tct tct ggt att ttc tgg gat aac gat agc tta 643
 Ala Pro Tyr Gln Asp Ser Ser Gly Ile Phe Trp Asp Asn Asp Ser Leu
 165 170 175

gct gct cta ctg gcg ttg gaa ctg aaa gct gat ctt ctg att ctt ctg 691
 Ala Ala Leu Leu Ala Leu Glu Leu Lys Ala Asp Leu Leu Ile Leu Leu
 180 185 190 195

agc gat gtt gaa ggt ctt tac aca ggc cct cca agt gat cct aac tca 739
 Ser Asp Val Glu Gly Leu Tyr Thr Gly Pro Pro Ser Asp Pro Asn Ser
 200 205 210

aag ttg atc cac act ttt gtt aaa gaa aaa cat caa gat gag att aca 787
 Lys Leu Ile His Thr Phe Val Lys Glu Lys His Gln Asp Glu Ile Thr
 215 220 225

ttc ggc gac aaa tca aga tta ggg aga ggg ggt atg act gca aaa gtc 835
 Phe Gly Asp Lys Ser Arg Leu Gly Arg Gly Gly Met Thr Ala Lys Val
 230 235 240

aaa gct gca gtc aat gca gct tat gct ggg att cct gtc atc ata acc 883
 Lys Ala Ala Val Asn Ala Ala Tyr Ala Gly Ile Pro Val Ile Ile Thr
 245 250 255

agt ggg tat tca gct gag aac ata gat aaa gtc ctc aga gga cta cgt 931
 Ser Gly Tyr Ser Ala Glu Asn Ile Asp Lys Val Leu Arg Gly Leu Arg
 260 265 270 275

gtt gga acc ttg ttt cat caa gat gct cgt tta tgg gct ccg atc aca 979
 Val Gly Thr Leu Phe His Gln Asp Ala Arg Leu Trp Ala Pro Ile Thr
 280 285 290

gat tct aat gct cgt gac atg gca gtt gct gcg agg gaa agt tcc aga 1027
 Asp Ser Asn Ala Arg Asp Met Ala Val Ala Ala Arg Glu Ser Ser Arg
 295 300 305

aag ctt cag gcc tta tct tcg gaa gac agg aaa aaa att ctg ctt gat 1075
 Lys Leu Gln Ala Leu Ser Ser Glu Asp Arg Lys Lys Ile Leu Leu Asp
 310 315 320

att gcc gat gcc ctt gaa gca aat gtt act aca atc aaa gct gag aat	1123
Ile Ala Asp Ala Leu Glu Ala Asn Val Thr Thr Ile Lys Ala Glu Asn	
325 330 335	
 gag tta gat gta gct tct gca caa gag gct ggg ttg gaa gag tca atg	1171
Glu Leu Asp Val Ala Ser Ala Gln Glu Ala Gly Leu Glu Glu Ser Met	
340 345 350 355	
 gtg gct cgc tta gtt atg aca cct gga aag atc tcg agc ctt gca gct	1219
Val Ala Arg Leu Val Met Thr Pro Gly Lys Ile Ser Ser Leu Ala Ala	
360 365 370	
 tca gtt cgt aag cta gct gat atg gaa gat cca atc ggc cgt gtt tta	1267
Ser Val Arg Lys Leu Ala Asp Met Glu Asp Pro Ile Gly Arg Val Leu	
375 380 385	
 aag aaa aca gag gtg gca gat ggt ctt gtc tta gag aag acc tca tca	1315
Lys Lys Thr Glu Val Ala Asp Gly Leu Val Leu Glu Lys Thr Ser Ser	
390 395 400	
 cca tta ggc gta ctt ctg att gtt ttt gaa tcc cga cct gat gca ctt	1363
Pro Leu Gly Val Leu Leu Ile Val Phe Glu Ser Arg Pro Asp Ala Leu	
405 410 415	
 gta cag ata gct tca ctt gcc atc cgt agt gga aat ggt ctt ctg ctg	1411
Val Gln Ile Ala Ser Leu Ala Ile Arg Ser Gly Asn Gly Leu Leu Leu	
420 425 430 435	

aag ggt gga aag gag gcc cgg cga tca aat gct atc tta cac aag gtg 1459
 Lys Gly Gly Lys Glu Ala Arg Arg Ser Asn Ala Ile Leu His Lys Val
 440 445 450

atc act gat gca att cca gag act gtt ggg ggt aaa ctc att gga ctt 1507
 Ile Thr Asp Ala Ile Pro Glu Thr Val Gly Gly Lys Leu Ile Gly Leu
 455 460 465

gtg act tca aga gaa gag att cct gat ttg ctt aag ctt gat gac gtt 1555
 Val Thr Ser Arg Glu Glu Ile Pro Asp Leu Leu Lys Leu Asp Asp Val
 470 475 480

atc gat ctt gtg atc cca aga gga agc aac aag ctt gtt act cag ata 1603
 Ile Asp Leu Val Ile Pro Arg Gly Ser Asn Lys Leu Val Thr Gln Ile
 485 490 495

aaa aat act aca aaa atc cct gtg cta ggt cat gct gat gga atc tgt 1651
 Lys Asn Thr Thr Lys Ile Pro Val Leu Gly His Ala Asp Gly Ile Cys
 500 505 510 515

cat gta tat gtc gac aag gct tgt gat acg gat atg gca aag cgc ata 1699
 His Val Tyr Val Asp Lys Ala Cys Asp Thr Asp Met Ala Lys Arg Ile
 520 525 530

gtt tct gat gca aag ttg gac tat cca gca gcc tgt aat gcg atg gaa 1747
 Val Ser Asp Ala Lys Leu Asp Tyr Pro Ala Ala Cys Asn Ala Met Glu
 535 540 545

acc ctt ctt gtg cat aag gat cta gag cag aat gct gtg ctt aat gag 1795
 Thr Leu Leu Val His Lys Asp Leu Glu Gln Asn Ala Val Leu Asn Glu
 550 555 560

ctt att ttt gct ctg cag agc aat gga gtc act ttg tat ggt gga cca 1843
 Leu Ile Phe Ala Leu Gln Ser Asn Gly Val Thr Leu Tyr Gly Gly Pro
 565 570 575

agg gca agt aag ata ctg aac ata cca gaa gca cgg tca ttc aac cat 1891
 Arg Ala Ser Lys Ile Leu Asn Ile Pro Glu Ala Arg Ser Phe Asn His
 580 585 590 595

gag tac tgt gcc aag gct tgc act gtt gaa gtt gta gaa gac gtt tat 1939
 Glu Tyr Cys Ala Lys Ala Cys Thr Val Glu Val Val Glu Asp Val Tyr
 600 605 610

ggt gct ata gat cac att cac cga cat ggg agt gca cac aca gac tgc 1987
 Gly Ala Ile Asp His Ile His Arg His Gly Ser Ala His Thr Asp Cys
 615 620 625

att gtg aca gag gat cac gaa gtt gca gag cta ttc ctt cgc caa gtg 2035
 Ile Val Thr Glu Asp His Glu Val Ala Glu Leu Phe Leu Arg Gln Val
 630 635 640

gat agc gct gct gtg ttc cac aac gcc agc aca aga ttc tca gat ggt 2083
 Asp Ser Ala Ala Val Phe His Asn Ala Ser Thr Arg Phe Ser Asp Gly
 645 650 655

ttc cga ttt gga ctt ggt gca gag gtg ggg gta agc acg ggc agg atc 2131
 Phe Arg Phe Gly Leu Gly Ala Glu Val Gly Val Ser Thr Gly Arg Ile
 660 665 670 675

cat gct cgt ggt cca gtc ggg gtc gaa gga tta ctt aca acg aga tgg 2179
 His Ala Arg Gly Pro Val Gly Val Glu Gly Leu Leu Thr Thr Arg Trp
 680 685 690

ata atg aga gga aaa gga caa gtt gtc gac gga gac aat gga att gtt 2227
 Ile Met Arg Gly Lys Gly Gln Val Val Asp Gly Asp Asn Gly Ile Val
 695 700 705

tac acc cat cag gac att ccc atc caa gct taaacaagac ttccgagtgt 2277
 Tyr Thr His Gln Asp Ile Pro Ile Gln Ala
 710 715

gtgtttgtgt atttggttga gacttgagga gagacacaga ggaggatggg cttttttgtt 2337

tctctctctgc ttagtactca tatctatca ttattattat tactactact tattattgaa 2397

accctcgctt atgtagtgtt ttgatttag ggtaggatt gcacacaaaa taagatccac 2457

tttaccactt agtcttgctc ataagtacga tgaagaacat ttaattagct tctctctctt 2517

tcattgtaag ctacctacac atttctgatc tttatcaaga tactactact tttc 2571

<210> 3
 <211> 1833
 <212> DNA
 <213> *Arabidopsis thaliana*
 <220>
 <221> CDS
 <222> 113...1612
 <301> Tomohiro Kiyasue, Yoshu Yoshiba, Kazuko
 Yamaguchi-Shinozaki, Kazuo Shinozaki
 <302> Title : A nuclear gene encoding mitochondrial
 proline dehydrogenase, an enzyme involved in
 proline metabolism, is upregulated by proline but
 downregulated by dehydration in *Arabidopsis*.
 <303> The Plant Cell
 <304> 8
 <306> 1323-1335
 <307> 1996-05-27
 <308> D83025
 <309> 1995-12-25
 <400> 3
 agcgtttaga aaaaaacgc gataaaacgc aaacatcaag caaacaaaa aaaaagagaa 60

gagaaattat tttttttgt tttcgttttc aaaaacaaaa tctttgaatt tt atg gca 118

Met Ala

1

acc cgt ctt ctc cga aca aac ttt atc cgg cga tct tac cgt tta ccc 166
 Thr Arg Leu Leu Arg Thr Asn Phe Ile Arg Arg Ser Tyr Arg Leu Pro

115	120	125	130	
gcc ggt gaa gat gcc gac gca gcc gct gag cgc gtg aga agc gtt tat				550
Ala Gly Glu Asp Ala Asp Ala Ala Ala Glu Arg Val Arg Ser Val Tyr				
	135	140	145	
gaa gct act ggt ctt aaa ggg atg ctt gtc tat ggc gtc gaa cac gcc				598
Glu Ala Thr Gly Leu Lys Gly Met Leu Val Tyr Gly Val Glu His Ala				
	150	155	160	
gat gac gct gta tct tgt gat gat aac atg caa caa ttc att cga acc				646
Asp Asp Ala Val Ser Cys Asp Asp Asn Met Gln Gln Phe Ile Arg Thr				
	165	170	175	
att gaa gct gcc aaa tct tta cca aca tct cac ttt agc tca gtg gtt				694
Ile Glu Ala Ala Lys Ser Leu Pro Thr Ser His Phe Ser Ser Val Val				
	180	185	190	
gtg aag ata act gcc att tgt cca att agt ctt ctg aaa cga gtg agc				742
Val Lys Ile Thr Ala Ile Cys Pro Ile Ser Leu Leu Lys Arg Val Ser				
	195	200	205	210
gat ctg ctg cgg tgg gaa tac aaa agt ccg aac ttc aaa ctc tca tgg				790
Asp Leu Leu Arg Trp Glu Tyr Lys Ser Pro Asn Phe Lys Leu Ser Trp				
	215	220	225	
aag ctc aaa tcg ttt ccg gtt ttc tcc gaa tcg agt cct ctc tac cac				838
Lys Leu Lys Ser Phe Pro Val Phe Ser Glu Ser Ser Pro Leu Tyr His				

230	235	240	
aca aac tca gaa ccg gaa ccg tta acc gcg gaa gaa gaa agg gag ctc			886
Thr Asn Ser Glu Pro Glu Pro Leu Thr Ala Glu Glu Glu Arg Glu Leu			
245	250	255	
gaa gca gct cat gga agg att caa gaa atc tgt agg aaa tgc caa gag			934
Glu Ala Ala His Gly Arg Ile Gln Glu Ile Cys Arg Lys Cys Gln Glu			
260	265	270	
tcc aat gta cca ttg ttg att gat gcg gaa gac aca atc ctc caa ccc			982
Ser Asn Val Pro Leu Leu Ile Asp Ala Glu Asp Thr Ile Leu Gln Pro			
275	280	285	290
gcg atc gat tac atg gct tat tca tcg gcg atc atg ttc aat gct gac			1030
Ala Ile Asp Tyr Met Ala Tyr Ser Ser Ala Ile Met Phe Asn Ala Asp			
295	300	305	
aaa gac cga cca atc gtt tac aac acg att cag gcg tac ttg aga gac			1078
Lys Asp Arg Pro Ile Val Tyr Asn Thr Ile Gln Ala Tyr Leu Arg Asp			
310	315	320	
gcc ggt gag aga ctg cat ttg gca gta caa aat gct gag aaa gag aat			1126
Ala Gly Glu Arg Leu His Leu Ala Val Gln Asn Ala Glu Lys Glu Asn			
325	330	335	
gtt cct atg ggg ttc aag ttg gtg aga ggg gct tac atg tct agc gaa			1174
Val Pro Met Gly Phe Lys Leu Val Arg Gly Ala Tyr Met Ser Ser Glu			

340	345	350	
cgt agc ttg gcg gat tcc ctg ggt tgc aag tcg cca gtc cac gac aca			1222
Arg Ser Leu Ala Asp Ser Leu Gly Cys Lys Ser Pro Val His Asp Thr			
355	360	365	370
att cag gat act cac tct tgt tac aat gat tgt atg aca ttc ctg atg			1270
Ile Gln Asp Thr His Ser Cys Tyr Asn Asp Cys Met Thr Phe Leu Met			
375	380		385
gag aaa gca tca aac ggt tct ggt ttc ggt gtc gtt ctc gca aca cat			1318
Glu Lys Ala Ser Asn Gly Ser Gly Phe Gly Val Val Leu Ala Thr His			
390	395		400
aac gct gat tcg ggg aga ctt gcg tcg agg aaa gcg agt gac ctc ggg			1366
Asn Ala Asp Ser Gly Arg Leu Ala Ser Arg Lys Ala Ser Asp Leu Gly			
405	410		415
atc gat aaa cag aac ggg aag ata gag ttt gca cag cta tat ggt atg			1414
Ile Asp Lys Gln Asn Gly Lys Ile Glu Phe Ala Gln Leu Tyr Gly Met			
420	425		430
tca gat gca ttg tcc ttc ggg tta aag aga gca ggg ttc aat gtt agc			1462
Ser Asp Ala Leu Ser Phe Gly Leu Lys Arg Ala Gly Phe Asn Val Ser			
435	440	445	450
aag tac atg ccg ttt gga ccc gtc gca acc gct ata ccg tat ctt ctc			1510
Lys Tyr Met Pro Phe Gly Pro Val Ala Thr Ala Ile Pro Tyr Leu Leu			

455	460	465	
cga cgc gct tat gag aac cgg gga atg atg gcc acc gga gct cat gac			1558
Arg Arg Ala Tyr Glu Asn Arg Gly Met Met Ala Thr Gly Ala His Asp			
470	475	480	
cgt caa ctc atg agg atg gaa ctt aag agg aga tta atc gcc ggg att			1606
Arg Gln Leu Met Arg Met Glu Leu Lys Arg Arg Leu Ile Ala Gly Ile			
485	490	495	
gcg taaagagaga gtatggagcc attaaatgaa attgggaaat gtagatgaat			1659
Ala			
aaatttcctc tatgtagttt aagaaattga aaacaaaaa ttataatata agaatggag			1719
taggtaagaa catttcctgt ggctaaatat tttcatgag ggactatgtt ttactatca			1779
atatatcatt cacaaatgta tattcacctt atcaataaaa atgcttttta cttt			1833